



PATENT ABSTRACTS OF JAPAN

(11) Publication number: **09212192 A**(43) Date of publication of application: **15 . 08 . 97**

(51) Int. Cl.
G10L 3/00
G10L 3/00
G06F 1/04

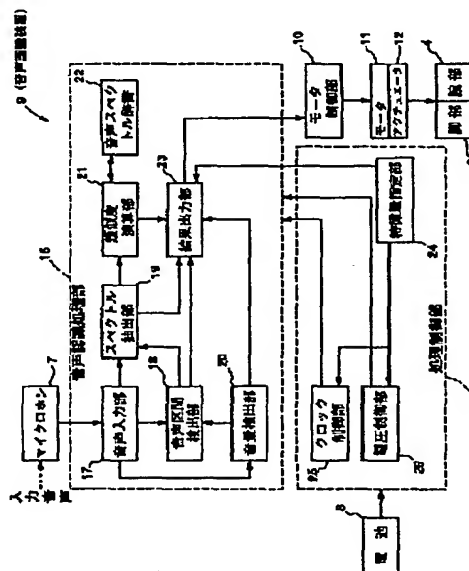
(21) Application number: **08045528**(71) Applicant: **RICOH CO LTD**(22) Date of filing: **07 . 02 . 96**(72) Inventor: **MIYAMOTO KEIICHI**(54) **SPEECH RECOGNITION DEVICE**

(57) Abstract:

PROBLEM TO BE SOLVED: To obtain necessary speech recognition precision and power consumption reduction by switching the kind of a speech recognition process according to recognition precision that an application requires and then performing the speech recognition process at a clock frequency corresponding to the throughput required by the actually used speech recognition process.

SOLUTION: Under specified recognition conditions, a process control part 16 is placed in operation to places only necessary parts of a speech recognition process part 15 selectively in operation at the clock frequency with the value of a source voltage needed to secure the specified speech recognition precision, thereby recognizing a speech inputted to a microphone 7. At the same time, a motor control part 10 is placed in operation according to this process result to perform operation corresponding to the contents of the speech.

COPYRIGHT: (C)1997,JPO



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平9-212192

(43) 公開日 平成9年(1997) 8月15日

(51) Int.Cl. ⁶	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L 3/00	5 7 1		G 1 0 L 3/00	5 7 1 J 5 7 1 K
	5 5 1			5 5 1 H
G 0 6 F 1/04	3 0 1		G 0 6 F 1/04	3 0 1 C

審査請求 未請求 請求項の数3 F D (全 7 頁)

(21) 出願番号 特願平8-45528

(22) 出願日 平成8年(1996) 2月7日

(71) 出願人 000006747

株式会社リコー

東京都大田区中馬込1丁目3番6号

(72) 発明者 宮本 恵一

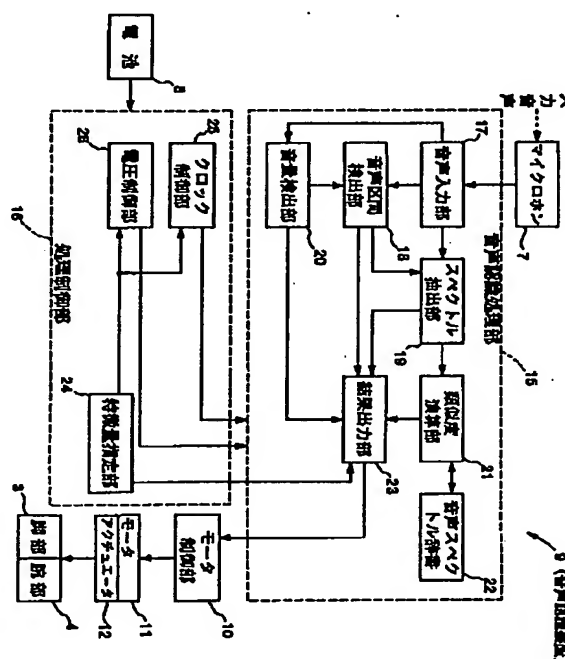
東京都大田区中馬込一丁目3番6号 株式会社リコー内

(54) 【発明の名称】 音声認識装置

(57) 【要約】

【課題】 本発明はアプリケーションが要求する認識精度に応じて、音声認識処理の種類を切り換え、これによって実際に使用される音声認識処理に必要な処理量に応じたクロック周波数で音声認識処理を行なわせて、必要な音声認識精度と、低消費電力化とを達成する。

【解決手段】 指定された認識条件で、処理制御部16を動作させて、指定された音声認識精度を確保するのに必要なクロック周波数、電源電圧の値で音声認識処理部15の必要な部分のみを選択的に動作させて、マイクロホン7に入力された音声を認識処理させるとともに、この処理結果に基づき、上記モータ制御部10を動作させて、上記音声の内容に対応する動作を行なわせる。



【従来の技術】音声認識を行なう音声認識装置は、音声の特徴を抽出する特徴抽出部と、この特徴抽出部で得られた特徴パターンがどの基準パターン（標準パターン）と類似しているかを見出す類似度演算部と、この類似度演算部で得られた各類似度に基づき、音声認識対象となっている音声かどの単語かを判定する判定部とを主要な構成要素にしていることが多い。この場合、上記特徴抽出部は、アナログ的には、バンドパスフィルタなどによって音声信号を構成する各周数成分の大きさなどの特徴情報を抽出して、特徴パターンなどを生成する方法を使用し、またデジタル的には、A/Dコンバータによって音声信号（アナログ信号）を音声データ（デジタル信号）に変換した後、DSP（デジタルシグナルプロセッサ）などによって上記音声データをデジタル的に処理して特徴パターンを生成する方法を使用することが多い。また、類似度演算部や判定部は、各種のデータ処理を行なうことができるマイクロプロセッサなどによって類似

【０００４】しかしながら、これら特開昭６２－２４５２９６号公報、特開平３－２０２８９９号公報に示す技術や特開昭５８－５５９９１号公報に示す技術では、次に述べるような問題があった。すなわち、特開昭６２－２４５２９６号公報や特開平３－２０２８９９号公報に示す技術では、音声が一定時間以上、入力されずに、電源が切られると、音声認識動作をさせたいとき、その都度、電源スイッチなどを操作しなければならず、操作が煩雑になってしまうという問題があった。また、特開昭５８－５５９９１号公報に示す技術では、一見、このような問題が無くなっているように見えるが、実際には、非常に簡単な積分回路によって、低消費電力の音声検出部を構成していることから、雑音が入力されたとき、これによつて音声検出部が誤動作してしまうという問題がある。さらに、音声が入力されてから、音声検出部によつて音声の入力が検知され、音声認識部で音声の認識処理が開始されるまで、ある程度の時間がかかってしまうことから、音声認識部が音声認識処理を開始したとき、音声の始端が欠落してしまい、音声認識精度が大幅に低

下してしまうという問題がある。そこで、BBD（バケッドブリゲード形素子）などのアナログ遅延素子を音声認識部の前段側に挿入し、音声検出部によって音声の入力を検出した後、音声認識部に音声信号公報が入力されるようにすることにより、上述した問題を解決することも考えられるが、このような方法では、もはや低消費電力を目指した構成とは言えなくなる。

【0005】さらに、このような従来の技術を実際の機器に適用すると、次に述べるような問題があった。すなわち、従来の技術を玩具などのアプリケーションに使用する場合、音声認識部の電源が切れている場合でも、何らかの方法を使用し、なるべく音声で動作させた方が望ましいことが多いが、上述した特開昭62-245296号公報や特開平3-202899号公報に示す技術では、これに対応することができない。さらに、上述した特開昭58-55991号公報に示す技術では、低消費電力で音声認識部を動作させているとき、通常の消費電力で音声認識部を動作させているときと同じ精度の高い認識精度を確保することができない。また、従来の技術を実際の機器などのアプリケーションに使用する場合、常に同じ様な認識結果が必要とは限らないことがある。例えば、クイズの答え（まる、ばつなど）を音声認識で回答させる場合、非常に高い認識率が要求されるため、このような要求に耐え得るだけの、認識特徴情報の抽出やマッチングアルゴリズムの計算が必要になる。またあるときには、認識結果に偶然性を取り入れた方が好ましいときがあり、このようなときには、比較的簡易な特徴情報やアルゴリズムで十分である。さらに、入力音声の量だけを判定して、「元気」、「意気消沈」などと判定するものであれば、もっと少ない処理量ですむ。このように、音声認識装置を搭載する機器によって、必要とされる認識結果の精度が異なるが、必要とされる認識精度に応じて、動的に動作を切り換える音声認識装置は未だ、開発されていない。

【0006】本発明は上記の事情に鑑み、請求項1では、アプリケーションが要求する認識精度に応じて、音声認識処理の種類を切り換え、これによって実際に使用される音声認識処理で必要な処理量に応じたクロック周波数で音声認識処理を行なわせて、必要な音声認識精度と、低消費電力化とを達成することができる音声認識装置を提供することを目的としている。請求項2では、アプリケーションが要求する認識精度に応じて、音声認識処理の種類を切り換え、これによって実際に使用される音声認識処理で必要な処理量に応じたクロック周波数、電源電圧で音声認識処理を行なわせて、必要な音声認識精度と、低消費電力化とを達成することができる音声認識装置を提供することを目的としている。請求項3では、アプリケーションが要求する認識精度に応じて、音声認識処理の種類を切り換え、これによって実際に使用される音声認識処理で必要な処理量に応じたクロック周

波数、電源電圧で音声認識処理を行なわせるとともに、認識処理動作に関与しない部分に対する給電を停止させて、必要な音声認識精度と、低消費電力化とを達成することができる音声認識装置を提供することを目的としている。

【0007】

【課題を解決するための手段】上記の目的を達成するために本発明による音声認識装置は、請求項1では、入力された音声音声信号公報に変換するマイクロホンと、このマイクロホンから出力される音声信号公報の特徴情報を相互に異なる抽出方法で抽出する第1～第N特徴抽出部と、これら第1～第N特徴抽出部に対して、クロック信号公報を供給して、指定された速度で動作させるクロック制御部と、音声認識の精度に応じて、特徴量を求め、この特徴量に応じて、上記クロック制御部から出力されるクロック信号公報の周波数を変化させる特徴量指定部とを備えたことを特徴としている。また、請求項2では、請求項1に記載の音声認識装置において、上記第1～第N特徴抽出部に対して、指定された電圧値を持つ電源電圧を供給する電圧制御部を備え、上記特徴量指定部で得られた特徴量に応じて、上記電圧制御部から出力される電源電圧の電圧値を変化させることを特徴としている。また、請求項3では、請求項1または2に記載の音声認識装置において、上記クロック制御部および上記電圧制御部は、上記第1～第N特徴抽出部のうち、上記特徴量指定部で得られた特徴量に対応し、特徴情報の抽出に関与しない部分に対するクロック信号公報の供給、電源電圧の供給を選択的に停止することを特徴としている。

【0008】上記の構成により、請求項1の音声認識装置では、音声認識の精度に応じて、特徴量指定部によって音声認識するのに必要な特徴量を求め、この特徴量に応じて、クロック制御部から出力されるクロック信号公報の周波数を変化させて、相互に異なる抽出方法で特徴情報を抽出する第1～第N特徴抽出部の動作速度を変化させて、マイクロホンから出力される音声データの特徴情報を抽出させることにより、アプリケーションにおいて、実際に使用される音声認識処理で必要な処理量に応じた動作速度で、音声認識処理を行なわせ、必要な音声認識精度と、低消費電力化とを達成する。また、請求項2では、請求項1に記載の音声認識装置において、上記特徴量指定部で得られた特徴量に応じて、電圧制御部から第1～第N特徴抽出部に供給される電源電圧の電圧値を変化させることにより、アプリケーションにおいて、実際に使用される音声認識処理で必要な処理量に応じた動作速度、電源電圧で、音声認識処理を行なわせ、必要な音声認識精度と、低消費電力化とを達成する。また、請求項3では、請求項1または2に記載の音声認識装置において、上記クロック制御部および上記電圧制御部によって、上記第1～第N特徴抽出部のうち、上記特徴量

指定部で得られた特徴量に対応し、特徴情報の抽出に関与しない部分に対するクロック信号公報の供給、電源電圧の供給を選択的に停止することにより、アプリケーションにおいて、実際に使用される音声認識処理で必要な処理量に応じた動作速度、電源電圧で、特徴情報の抽出に関与する部分のみを動作させて、音声認識処理を行なわせ、必要な音声認識精度と、低消費電力化とを達成する。

【0009】

【発明の実施の形態例】以下、本発明を図面に示した形態例に基づいて詳細に説明する。図1は本発明による音声認識装置の一形態例を使用したロボット玩具の一例を示す正面図である。この図に示すロボット玩具1は、円柱状に形成される胴体部2と、この胴体部2の下部に可動自在に取り付けられる2本の脚部3と、上記胴体部2の上側部に可動自在に取り付けられる2本の腕部4と、上記胴体部2の上部に固定される首部5と、この首部5上に固定される頭部6と、この頭部6の側部に固定されるマイクロホン7と、上記胴体部2内に内蔵される電池8と、上記胴体部2内に内蔵され、上記電池8によって得られる電力により、上記マイクロホン7に入力された

40 音声認識処理する音声認識装置9と、上記胴体部2内に内蔵され、上記音声認識装置9から出力される制御信号公報（駆動信号公報）に応じた駆動電圧を生成するモータ制御部10と、上記胴体部2内に内蔵され、上記モータ制御部10から出力される駆動電圧に応じた駆動力を発生する複数のモータ11と、上記胴体部2内のうち、上記各脚部3、各腕部4の付け根部分に内蔵され、上記各モータ11によって得られた駆動力により、脚部3を動かしたり、腕部4を動かしたりする複数のアクチュエータ12とを備えている。

【0010】そして、ロボット玩具1の使用者が予め設定されている単語のうちのいずれかを発声したとき、音声認識処理に必要な部分のみに電源を供給しながら、上記音声音声認識して、この音声で指定された動作を行なう。上記音声認識装置9は、図2に示す如く音声認識処理部15と、処理制御部16とを備えており、指定された認識条件で、処理制御部16を動作させて、指定された音声認識精度を確保するのに必要なクロック周波数、電源電圧の値で、音声認識処理部15の必要な部分のみを選択的に動作させて、マイクロホン7に入力された音声認識処理させるとともに、この処理結果に基づき、上記モータ制御部10を動作させて、上記音声の内容に対応する動作を行なわせる。音声認識処理部15は、音声入力部17と、音声区間検出部18と、スペクトル抽出部19と、音量検出部20と、類似度演算部21と、音声スペクトル辞書22と、結果出力部23とを備えており、上記処理制御部16から供給されるクロック信号公報と、電源電圧とに基づき、マイクロホン7から出力される音声信号公報をデジタル化して音声データにした

後、この音声データに基づき、音声区間情報を抽出しながら、スペクトルによる音声認識処理が指定されていれば、上記音声区間情報に基づき、上記音声データのスペクトルを抽出して、音声スペクトル辞書22を用いた類似度演算処理により、上記音声データで示される単語を特定し、また上記音声区間情報による音声認識処理が指定されていれば、上記音声データの音声区間情報に基づき、上記音声データで示される単語を特定し、また音量情報による音声認識処理が指定されていれば、上記音声データの音量情報に基づき、上記音声データで示される単語を特定した後、特定された単語に対応する駆動信号公報を生成し、これを上記モータ制御部10に供給する。

【0011】上記音声入力部17は、上記マイクロホン7から出力される音声信号公報を増幅する増幅回路、この増幅回路から出力される音声信号公報をフィルタリングしてノイズなどを取り除くフィルタ回路、このフィルタ回路から出力される音声信号公報を音声データに変換するA/D変換回路などを備えており、上記マイクロホン7から出力される音声信号公報を取り込んで、これを増幅するとともに、フィルタリングして、ノイズを取り除いた後、A/D変換して音声データにし、これを上記音量検出部20と、上記音声区間検出部18と、スペクトル抽出部19とに供給する。音量検出部20は、上記音声入力部17から出力される音声データの音量を検出して、この音声データの音量を示す音量情報を生成し、これを上記音声区間検出部18と、結果出力部23とに供給する。

【0012】音声区間検出部18は、上記音声入力部17の出力と、上記音量検出部20の出力とに基づき、上記マイクロホン7に音声が入力されたとき、これを検知して音声検知情報を生成するとともに、この音声データ中の無音区間、例えば音声と音声との境や「ストップ」などの「スト」と「ブ」との間にある「ッ」などの促音などを検出して、これらの検出結果を無音区間情報とし、これら音声検知情報、無音区間情報をスペクトル抽出部19と、結果出力部23とに供給する。スペクトル抽出部19は、上記無音区間検出部18から出力される音声検知情報、無音区間情報に基づき、上記音声入力部17から出力される音声データを音声区間毎に分けて周波数分析し、タイムスペクトラムパターン、LPC係数、零クロス数、パワー値などの特徴情報を抽出した後、これらの特徴情報に基づき、上記音声データの特徴を示す特徴パターンを生成し、これを類似度演算部21と、結果出力部23とに供給する。

【0013】また、音声スペクトル辞書22は、認識対象となる単語、例えば「めざめよ」、「バスター」、「進め」、「やっつけろ」などの特徴を示す標準パターンが格納されており、上記類似度演算部21から読み出し指示が出されたとき、指定された単語の標準パターン

を読み出し、これを上記類似度演算部21に供給する。類似度演算部21は、上記スペクトル抽出部19から上記音声データの特徴を示す特徴パターンが出力されたとき、この特徴パターンと、上記音声スペクトル辞書22に格納されている各標準パターンとをパターンマッチング処理して、これらの類似度を演算し、この演算処理によって、上記特徴パターンに対して、最も距離が小さく、かつ予め設定されている基準値より小さい距離を持つ標準パターンがあれば、これを上記音声データに対応する単語と判定し、この判定結果を結果出力部23に供給する。

【0014】結果出力部23は、上記処理制御部16から出力される音声認識指示に基づき、上記音量検出部20、音声区間検出部18、スペクトル抽出部19、類似度演算部21、音声スペクトル辞書22のうち、必要な部分を動作状態にして、音声認識処理を行なわせるとともに、この音声認識処理で得られた認識結果を取り込んで、上記音声データに対応する単語を決定し、この決定結果に応じた駆動信号公報を生成し、これを上記モータ制御部10に供給する。

【0015】上記処理制御部16から出力される音声認識指示によってスペクトル抽出処理を使用した音声認識が指定されていれば、上記音量検出部20、音声区間検出部18、スペクトル抽出部19、類似度演算部21、音声スペクトル辞書22の全てが動作状態にされて、図3に示す如く上記音声データが「めざめよ」、「バスター」、「進め」、「やっつけろ」のいずれであるか判定され、また上記音声認識指示によって無音区間数を使用した音声認識が指定されていれば、上記音量検出部20、音声区間検出部18、スペクトル抽出部19、類似度演算部21、音声スペクトル辞書22のうち、上記音量検出部20、音声区間検出部18のみが動作状態にされて、上記音声データが「ター」、「パッパッパ」のいずれであるか判定される。また、上記音声認識指示によって音量を使用した音声認識が指定されていれば、上記音量検出部20、音声区間検出部18、スペクトル抽出部19、類似度演算部21、音声スペクトル辞書22のうち、上記音量検出部20のみが動作状態にされて、上記音声データが「大きな声」、「小さな声」のいずれであるか判定される。

【0016】また、処理制御部16は、ロボット玩具1の使用者によって指定された音声認識の指示に基づき、指定された音声認識指示を出すとともに、この音声認識指示に応じてクロック周波数指示信号公報と電圧値指示信号公報とを生成する特徴量指定部24と、この特徴量指定部24から出力されるクロック周波数指示信号公報に応じた周波数のクロック信号公報を生成するクロック制御部25と、上記特徴量指定部24から出力される電圧値指示信号公報に応じた電圧値を持つ電源電圧を生成する電圧制御部26とを備えている。ロボット玩具1の

使用者によってスペクトル抽出を用いた音声認識が指示されているときには、上記音声認識処理部15に対して、スペクトル抽出を用いて音声認識処理を行なうことを示す音声認識指示を出すとともに、最も高い周波数を持つクロック信号公報、最も高い電圧値を持つ電源電圧を生成し、これを上記音声認識処理部15に供給する。

【0017】また、ロボット玩具1の使用者によって無音区間を用いた音声認識が指示されているときには、上記音声認識処理部15に対して、無音区間を用いて音声認識処理を行なうことを示す音声認識指示を出すとともに、通常の周波数を持つクロック信号公報、通常の電圧値を持つ電源電圧を生成し、これを上記音声認識処理部15に供給する。これによって、音声認識装置9を構成している各素子、例えばCMOS回路など、クロック信号公報が切り替わる際に、電力を消費する素子の消費電力が低く抑さえられるとともに、各素子で使用される電力（電流×電圧）が低く抑さえられて、消費電力の低減が図られる。また、ロボット玩具1の使用者によって音量区間を用いた音声認識が指示されているときには、上記音声認識処理部15に対して、音量を用いて音声認識処理を行なうことを示す音声認識指示を出すとともに、最も低い周波数を持つクロック信号公報、最も低い電圧値を持つ電源電圧を生成し、これを上記音声認識処理部15に供給する。これによって、音声認識装置9を構成している各素子、例えばCMOS回路など、クロック信号公報が切り替わる際に、電力を消費する素子の消費電力が低く抑さえられるとともに、各素子で使用される電力（電流×電圧）が低く抑さえられて、さらに消費電力の低減が図られる。

【0018】このように、この形態例では、指定された認識条件で、処理制御部16を動作させて、指定された音声認識精度を確保するのに必要なクロック周波数、電源電圧の値で音声認識処理部15の必要な部分のみを選択的に動作させて、マイクロホン7に入力された音声認識処理させるとともに、この処理結果に基づき、上記モータ制御部10を動作させて、上記音声の内容に対応する動作を行なわせるようにしたので、アプリケーションが要求する認識精度に応じて、音声認識処理の種類を切り換え、これによって実際に使用される音声認識処理で必要な処理量に応じたクロック周波数、電源電圧で音声認識処理を行なわせるとともに、認識処理動作に関与しない部分に対する給電を停止させ、これによって必要な音声認識精度と、低消費電力化とを達成することができ。

【0019】また、上述した形態例においては、音声認識処理の種類に応じて、クロック周波数、電源電圧、電源の供給対象となる回路とを全て切り換えるようにしているが、これらのうちのいずれか、例えばクロック周波数のみを切り換えたり、クロック周波数および電源電圧のみを切り換えたりするようにしても、上述した形態例

とはほぼ同様な効果を得ることができる。また、上述した形態例においては、ロボット玩具1に対して、音声認識装置9を装着するようにしているが、携帯性を重視した電子機器、例えば電子手帳、ハンディターミナルなどに搭載するようにしても良い。このような電子機器に音声認識装置9を搭載すれば、必要な音声認識の精度を確保するのに必要な最も低い消費電力で、音声認識を行なうことができ、これによって電池の消耗を最低にしながら、長い時間、稼動させることができる。

【0020】

【発明の効果】以上説明したように本発明によれば、請求項1では、アプリケーションが要求する認識精度に応じて、音声認識処理の種類を切り換え、これによって実際に使用される音声認識処理で必要な処理量に応じたクロック周波数で音声認識処理を行なわせて、必要な音声認識精度と、低消費電力化とを達成することができる。請求項2では、アプリケーションが要求する認識精度に応じて、音声認識処理の種類を切り換え、これによって実際に使用される音声認識処理で必要な処理量に応じたクロック周波数、電源電圧で音声認識処理を行なわせて、必要な音声認識精度と、低消費電力化とを達成することができる。請求項3では、アプリケーションが要求する認識精度に応じて、音声認識処理の種類を切り換え、これによって実際に使用される音声認識処理で必要*

*な処理量に応じたクロック周波数、電源電圧で音声認識処理を行なわせるとともに、認識処理動作に関与しない部分に対する給電を停止させて、必要な音声認識精度と、低消費電力化とを達成することができる。

【図面の簡単な説明】

【図1】本発明による音声認識装置の一形態例を使用したロボット玩具の一例を示す正面図である。

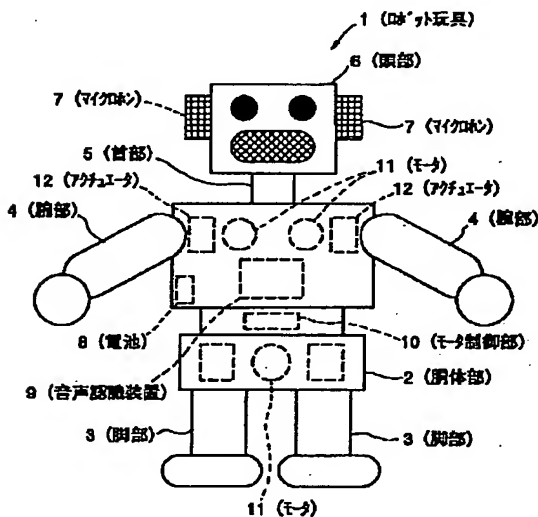
【図2】図1に示す音声認識装置の詳細な回路構成例を示すブロック図である。

10 【図3】図2に示す音声認識装置の処理例を示す表である。

【符号の説明】

1…ロボット玩具、2…胴体部、3…脚部、4…腕部、5…首部、6…頭部、7…マイクロホン、8…電池、9…音声認識装置、10…モータ制御部、11…モータ、12…アクチュエータ、15…音声認識処理部、16…処理制御部、17…音声入力部、18…音声区間検出部（第1～第N特徴抽出部）、19…スペクトル抽出部（第1～第N特徴抽出部）、20…音量検出部（第1～第N特徴抽出部）、21…類似度演算部（第1～第N特徴抽出部）、22…音声スペクトル辞書（第1～第N特徴抽出部）、23…結果出力部、24…特徴量指定部、25…クロック制御部、26…電圧制御部

【図1】



【図3】

単語番号	種別	単語音声	動作
1	スペクトル識別	めざめよ	ガオーと音を出し腹を上げる
2	スペクトル識別	バスター	全身をがたがた揺わせる
3	スペクトル識別	進め	足を動かして前進する
4	スペクトル識別	やっつけろ	腕を振る
5	無音区間数判別	ター	前へ進む
6	無音区間数判別	パッパッパ	後ろへ進む
7	無音判別	(大きな声)	動きが激しくなる
8	無音判別	(小さな声)	動きがおとなしくなる

(7)

【図2】

